UNCLASSIFIED

AD 406 910

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION
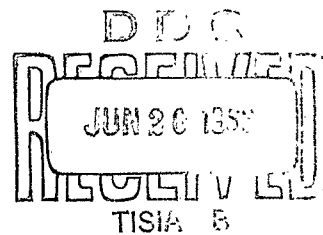
CAMERON STATION, ALEXANDRIA, VIRGINIA

UNCLASSIFIED

63-4-1

SP-986/000/0

Stimulus Sampling Pattern Recognitio

15 March 196

SP-986/000/01

Stimulus Sampling Pattern Recognition

Frank N. Marzocco

March 15, 1963

A-1162

ABSTRACT

A generalized computer model has been constructed to simulate the be-
havior of organisms according to various versions of stimulus sampling
theory. Results already reported dealt with the learning behavior of rats
in a T-maze. With the addition of new routines to specify the simulated
environment of the simulated organism, a similar basic program has learned
to discriminate among visual patterns consisting of carelessly hand-drawn
alphabetic characters, each of which is presented in a number of variations.
Input to the program is a binary representation of the presence or absence
of parts of the figure in a 20 x 20 matrix superimposed on the figure to
be recognized. Each cell of the matrix is treated as a separate stimulus
having various numbers of elements associated with all the responses that
can be made. A random sample of elements from stimuli present in the visual
field is used to determine the response on any trial. When a correct re-
sponse is made, all elements sampled on that trial become associated with
that response. Changes in learning rates occur as the probability of se-
lecting available stimulus elements is changed.

(

BACKGROUND

Mathematical models intended to predict or explain the behavior of living organisms frequently make use of the stimulus sampling principle. The situation in which behavior occurs is analyzed into stimulus classes, and the behavior itself is analyzed into classes of response. Interpretations for the model are observable stimuli and responses; predictions are ordinarily in terms of response probabilities, given the conditions that affect stimulus configurations. In deriving properties for these models, stimuli may be treated as if they were composed of a number of elements or cues, each of which is associated with one of the possible responses. The response that occurs at any moment is determined by a sample from the totality of stimulus elements available at the time. Each cue sampled on a given trial provides a tendency for occurrence of the associated response, and learning occurs as a change in the associations. The amount and kind of change is a function of the particular response that takes place in a given situation. Many variations are possible within the stimulus sampling framework, and a number of different models have been introduced since the original work by Estes (1). A recent review (2) provides background and an excellent description of current research being done on models of this type.

While the stimulus sampling principle has proved a powerful tool for understanding behavior in relatively simple laboratory situations, the models have been difficult to apply to the description of complex events. There are problems both in making some of the computations and in establishing correspondence between the model and observations on the real world. It is possible, however, to avoid some of these problems by employing the technique of computer simulation. The computer is programmed to represent a set of basic behavior processes which accord with a stimulus sampling model, and the processes are combined in ways that represent their multiple occurrences in the situation being simulated. Not only should this approach permit fairly easy extension of the model's domain of interpretation; it also provides a tool that may be helpful in understanding details of individual behavior.

A computer program that simulated behavior of rats in a T-maze (3) demonstrated feasibility of the approach. That program was capable of handling several sources of stimulation. However, in applying the program to some problems of discrimination learning, it became evident that the method of processing data was extremely inefficient and that the program was not very flexible in use. Both a new processing scheme and a number of new input-output routines were therefore devised. The problem of visual character recognition was chosen to test the machine because the problem was intrinsically interesting and would test many of the machine's new capabilities.

DESIGN OF THE PROGRAM

Stimuli in the computer program, as in mathematically expressed stimulus sampling models, correspond to particular event classes as an external observer would identify them in the environment of the organism whose behavior is under study. Names for the stimuli to be observed are provided on punched cards or magnetic tape. From these, the program sets up a table used both for internal processing and for naming the stimuli in outputs. Similarly, responses of the computer program correspond to the occurrence of particular event classes as an external observer would identify them in the behavior of the organism under study. Names of the responses to be observed in a set of runs are also typically provided on punched cards or magnetic tape; the program sets up another table for these. In the character recognition problem, stimuli correspond to discrete areas in a visual field. The areas are numbered sequentially and the numbers serve as stimulus names. Response names are not provided in advance for this application, however. Instead the program is given a set of blanks to construct the response table, and the blanks are filled in whenever a new pattern name is encountered.

Prior to any set of runs, the computer must be provided with an association state. A new state may be established from card or tape data, or a final state preserved on tape from an earlier run may be used as a new initial state. The treatment of associations differs significantly from that in the T-maze simulation program. Rather than identify separately each stimulus element under consideration, the new program tabulates the number of elements from each stimulus associated at any moment with each response and maintains these counts in a packed array. An array for the pattern recognition program could thus be regarded as an $m \times n$ matrix, where m is the number of areas into which the visual field is divided and n is the maximum number of character names that the program must learn to handle.

At the beginning of each simulated trial, a special subroutine indicates which of the possible stimuli are present and which of the possible responses may occur. The main program behaves as if it had taken a sample of elements from each stimulus present in the situation to determine the response on that trial. Repeated sampling from the pool of elements corresponding to a particular stimulus-response combination would produce a binomial sample size distribution with mean Np and variance $Np(1-p)$, where N is the total number of elements in the pool and p is the probability that any element is selected. With large numbers of elements, a good approximation to the theoretically exact binomial distribution is provided by a normal distribution with the same mean and variance. In actual operation, therefore, the program determines the number of elements to select by choosing for each combination a single pseudorandom value from the normal distribution with proper mean and variance. The computation is performed only for combinations in which the stimulus is present and the response possible. To decide which response occurs on the trial; the number of sampled elements associated with each response is

counted and the response with the largest count is emitted. In the case of
character recognition, a stimulus is considered present on any trial if any
part of the character presented on that trial appears in the corresponding
area of the visual field. Available responses ordinarily include the names
of all patterns that may be presented, although an option is included where-
by the machine may be forced to make a correct response.

Another special subroutine determines whether a response is to be reinforced.
For character recognition, the response emitted on any trial is compared with
the name of the pattern presented on that trial. If these are different, no
change is made in the associations. When a correct response has been produced,
however, the entries in the packed array are adjusted so that all stimulus
elements sampled on that trial become associated with the response that oc-
curred.

Alphabetic characters originally prepared for use with another pattern recog-
nition program (4, 5) were already available on punched cards and were there-
fore used as data inputs for the new program. The patterns, one of which is
shown in Figure 1, were prepared by sketching the characters inside a 20 x 20
grid. Any cell of the grid in which a part of the pattern appeared was coded
with a one and the remaining cells left blank. The ordered set of 400 ones
and blanks, together with the pattern name and other identifying information,
were keypunched according to a standard packing format. Each set of these
cards describes a single token of a given character. Input of any set causes
the computer to generate and store a table which lists, as stimuli present for
that token, the cells coded one in the input. The table also lists the cor-
rect name of the pattern; that name is added to the list of possible responses
if it does not already appear there.

A single input tape used for the runs described below named 400 stimuli and
reserved space for them. It also provided space for 26 responses, although
only four tokens of each pattern A, B, and C were included on the tape and
only these three response names were used. All responses were considered
equally likely at the beginning of the run, and the tape specified 100,000 cues
from each stimulus initially associated with each possible response. Control
cards provided for each run determined the amount of detail desired in the out-
put and caused the input tape to be preserved without change from run to run.
The order of presentation was also fixed. Additional input parameters set
the probability of sampling any element from stimuli present on a given trial,
determined the number of passes through the 12 pattern tokens, and also deter-
mined for each pass whether the simulated organism was free to choose among
the three responses on any trial or forced to make the correct response. In
the runs described below, the correct response was always forced for the first
pass through the token set, and the simulated organism was permitted to re-
spond freely on succeeding passes. Whether the pass was free or forced, cor-
rect responses were always reinforced and incorrect responses never reinforced.

Special subroutines required to perform pattern recognition were written in
JOVIAL. General processing routines were written in TAC. All runs were per-
formed on a Philco 2000 computer.

RESULTS AND DISCUSSION

When the parameter that controls probability of sampling any available stimu-
lus element was set at .001, the program produced the result shown in Table 1.
The proportion of successes obtained on the first free-response pass was .458,
while the expected proportion under the hypothesis of random choice with in-
dependent alternatives is .333. The critical ratio for the difference is
2.25, and the probability that a ratio this large would occur by chance is
less than 3%. Thus it seems reasonable to conclude that some learning took
place on the single preceding trial in which the correct response was forced
for each presentation of a pattern token. It is also apparent from inspec-
tion of the table that learning continued to take place over the 12 trials,
although there is considerable variability from one trial to another and none
of the simulated subjects learned to discriminate perfectly among all the
characters.

Other sets of runs demonstrate that the machine is capable of mastering at
least the limited task of discriminating among the 12 input characters. The
only change in procedure was to vary the probability of selecting any availa-
ble cue. For one set of runs the parameter was set at .005. Of the six sub-
jects run under this condition, one mistakenly responsed A for a B pattern,
one mistakenly responded B for a token of C, and one mistakenly responded B
for a different token of C. All three errors occurred on the first free-re-
sponse pass for these subjects; the correct response was emitted on each of
the remaining 861 free-response trials. Completely error-free runs occurred
on every pass when the values .01, .02, and .05 were used for the sampling
parameter. The machine broke down, however, when larger values were tried.
Each of the runs for which the sampling probability was set at .10, .15, or
.20 terminated prematurely because the number of associations for some sti-
mulus-response combination always fell too low for application of the sampling
technique used in the experiment. This feature indicates that it is possible
to administer too strong a reinforcement even with the use of other sampling
methods, since emission of a correct response to a given stimulus may remove
too many of the cues required to form other responses also correct for that
stimulus.

Tests performed on the machine so far have been designed more to expose pos-
sible faults in the programming than to provide useful data. Nevertheless
several conclusions can be drawn from the data. The computer with a stimulus
sampling program can improve in performance when presented the task of dis-
criminating among a limited set of patterns, and the rate of improvement is
sensitive to the sampling probability parameter. The program is capable of
mastering the task, and such mastery can be achieved quite rapidly.

Whether the stimulus sampling technique can be used to design machines that perform effective pattern recognition against a wide range of inputs has not been established. However, features of the computer with a stimulus sampling program resemble some which have already been included in pattern recognition devices that perform well--an outstanding example is the machine reported by Baran and Estrin (6). This fact, together with the speed with which simulated subjects under several conditions achieved apparently perfect discrimination, justifies further examination of the stimulus sampling principle as a source of ideas for pattern recognition machines.

Whether the machine interpretation of stimulus sampling theory can produce valid models for discrimination and pattern recognition by living organisms has also not been established. Trial-by-trial behavior of the simulated subjects with the smallest sampling probability might have been due in large part to the particular patterns and method of presentation; however, the data show many of the characteristics that would be expected from living subjects. Mathematically expressed stimulus sampling models have been used successfully to study a number of special discrimination learning problems; many of these studies are discussed by Estes (2, 7). The computer simulation method provides at least an additional tool--and perhaps a unique one.

REFERENCES

1. Estes, W. K. Toward a statistical theory of learning. Psychol. Rev.,
      1950, 57, 95-107.

2. Estes, W. K. Learning theory. Ann. Rev. Psychol., 1962, 13, 107-144.

3. Marzocco, F. N., and Bartram, P. R. Statistical learning models for be-
      havior of an artificial organism. In E. E. Bernard and M. R. Kare (Eds.),
      Biological prototypes and synthetic systems, vol. 1, New York: Plenum
      Press, 1962. Pp. 88-96.

4. Uhr, L., and Vossler, C. A pattern recognition program that generates,
      evaluates, and adjusts its own operators. Proc. West. Joint Computer
      Conf., 1961, 19, 555-569.

5. Vossler, C., and Uhr, L. A computer simulation of pattern perception
      and concept formation. In E. E. Bernard and M. R. Kare (Eds.), Bio-
      logical prototypes and synthetic systems, vol. 1, New York: Plenum
      Press, 1962. Pp. 233-243.

6. Baran, P., and Estrin, G. An adaptive character reader. IRE WESCON
      Conv. Rec., 1960, Part 4, 29-41.

7. Estes, W. K. The statistical approach to learning theory. In S. Koch
      (Ed.) Psychology: a study of a science, vol. 2, General systematic
      formulations, learning, and special processes. New York: McGraw-Hill,
      1959. Pp. 380-491.

Table 1

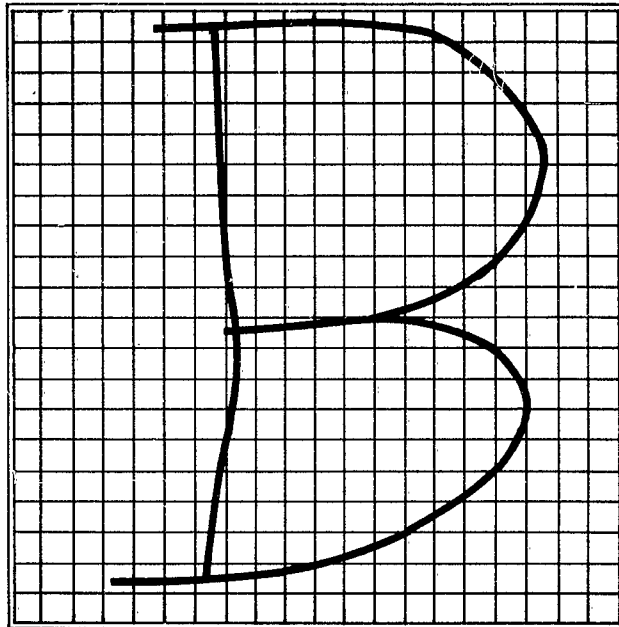Correct Responses on Free-Response Passes, Sampling Probability .001

Subject	Pass											
	1	2	3	4	5	6	7	8	9	10	11	12
1	4	4	7	2	8	10	7	10	8	10	8	9
2	7	5	4	8	5	9	8	8	7	8	9	8
3	6	7	6	8	8	8	9	7	10	9	9	9
4	5	6	7	6	5	7	7	6	9	9	9	9
5	5	7	9	8	7	8	8	8	10	8	11	10
6	6	7	8	5	7	8	8	9	9	11	10	8
Mean	5.5	6.0	6.8	6.2	6.7	8.3	7.8	8.0	8.8	9.2	9.3	8.8

ORIGINAL PATTERN

CODED PATTERN
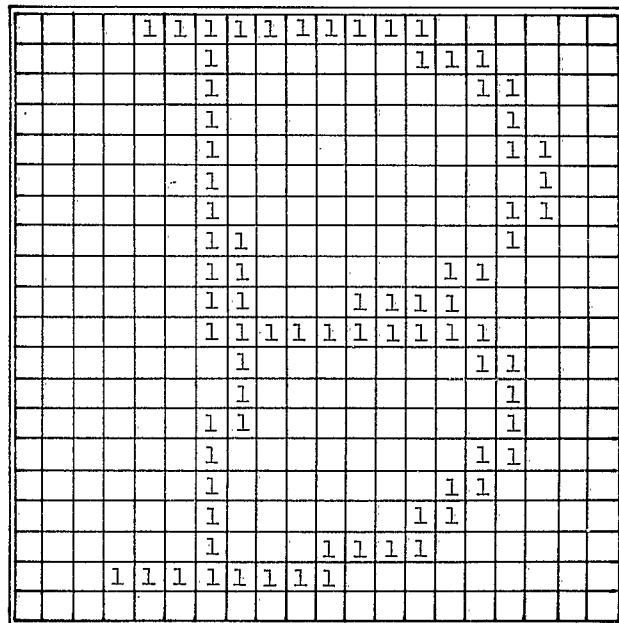
Figure 1.

DESCRIPTORS: Pattern Recognition.

Reports that a generalized computer
model has been constructed to
simulate the behavior of organisms
according to various versions of
stimulus-sampling theory. Results

already reported dealt with the
learning behavior of rats in a T-
maze. With the addition of new
routines to specify the simulated
environment of the simulated organism,
a similar basic program has learned
to discriminate among visual patterns
consisting of carelessly hand-drawn
alphabetic characters, each of
which is presented in a number of
variations. Input to the program
is a binary representation of the
presence or absence of parts of the
figure. Changes in learning rates
occur as the probability of selecting
available stimulus elements is
changed.